DOCUMENT RESUME

ED 031 173 HE 001 055

By-Beanblossom, Gary F.
What Do the CLEP General Examinations Measure?
Washington Univ., Seattle. Bureau of Testing.
Pub Date Jun 69
Note-11p.
EDRS Price MF-$0.25 HC-$0.65
Descriptors-*Aptitude Tests, *Cognitive Development, *Factor Analysis, Higher Education, Liberal Arts, *Predictive Ability (Testing), *Student Testing, Test Validity

Factor analysis was used to test the hypothesis that College Level Examination Program (CLEP) General Examinations, when administered to students who have completed 2 years of college, do not measure anything different from that measured by the traditional battery of pre-college aptitude examinations. The CLEP examinations were taken in the fall of 1968 by 333 University of Washington juniors, all but 2 of whom had previously taken the Washington Pre-College Test (WPC). Three CLEP scores (in the areas of social science-history, natural science, and humanities) and 11 WPC scores were intercorrelated and factor analyzed --using a computer program described by Kaiser in 1959-- to reveal whether CLEP scores increase the factorial complexity of the WPC battery. Study results revealed that while the CLEP General Examinations are valid indicators of educational achievement, they do not measure anything uniquely different from that measured by the WPC battery and should be used with caution for evaluating liberal arts curricula. Students who score well on pre-college tests may achieve better in the liberal arts areas covered by the CLEP examinations, since they may be more intellectually curious and highly motivated. However, the CLEP natural science examination scores for this sample of juniors is just as predictable as scores for vocabulary or reading comprehension tests taken 3 years earlier. (WM)

N-Y

ED031173 HE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Bureau of Testing

University of Washington

June 1969

What Do the CLEP General Examinations Measure?

Gary F. Beanblossom

Eleven scores from the Washington Pre-College battery along with scores in the three CLEP General Examination areas (social science-history, natural science, and humanities) were factor analyzed to reveal whether CLEP scores, ostensibly tapping college liberal arts achievement, measure something apart from that measured by the traditional pre-college battery of aptitude and achievement tests. Four factors were identified: (1) quantitative, (2) verbal, (3) English mechanics, and (4) masculine interest. Social science-history and humanities load almost exclusively on the verbal factor, suggesting that reading comprehension and vocabulary are the critical skills necessary to do well on these tests. Natural science also shows a substantial verbal component, but also loads rather well on masculine interest and moderately well on the quantitative factor. These results would indicate that the use of CLEP exams to evaluate liberal arts curricula should be accompanied by a proper amount of cautious skepticism regarding the validational claims of the test publishers.

Background. Three of the five College Level Examination Program (CLEP) General Examinations were administered to 333 University juniors during Autumn Quarter 1968 to evaluate student proficiency in lower division liberal arts areas and establish normative data whereby certain portions of lower division course requirements not completed might be waived. Results from this study have recently been reported (Beanblossom, 1969). Exceptionally high correlations between the three CLEP scores (social science-history, natural science, and humanities) were found, even though, on the surface,

---

Bureau of Testing Project 0668-102

HE 001 855

one would expect rather low correlations since students with many credits in natural science (and hence presumably higher natural science scores) are apt to have few credits in social science and humanities (and hence presumably lower social science-history and humanitites CLEP scores). This led to the suspicion that the exams were "measuring a common underlying factor, perhaps reading speed and comprehension or some facet of general academic competence." This report attempts to test this suspicion through the use of factor analytic procedures.

What CLEP Ostensibly Measures. General examinations in mathematics and English composition are offered in addition to the three exams that roughly parallel the conventional liberal arts areas of social science, natural science, and humanities. The literature (CLEP, 1968, p. 7) informs us that "the General Examinations are intended to provide a comprehensive measure of undergraduate achievement in the five basic areas of the liberal arts listed above." Furthermore the exams attempt " . . . . to assess a student's knowledge of fundamental facts and concepts, his ability to perceive relationships, and his understanding of the basic principles of the subject." (Emphasis is mine). In another publication (CLEP, 1967, p. 5) we are told that the exams stress understanding of facts, "not merely retention." But if the exams stress "knowledge" and "understanding" of facts, "retention" would seemingly be a necessary precondition; but perhaps it is more a matter of emphasis. At any rate, that elusive thing called educational "achievement," as defined here, is a multi-dimensional concept involving understanding and perception of relationships as well as substantive knowledge of liberal arts areas. But does CLEP, administered to students who

have completed two years of college, really measure something apart from that measured by the traditional battery of "aptitude" exams that students commonly undergo prior to college?

The Washington Pre-College Test. All but two of the 333 students taking CLEP had previously taken the Washington Pre-College Test (WPC) before entering the University. The majority completed the WPC battery soon after the beginning of their senior year in high school. A brief description of the tests (WPC, 1967-68, pp. 13-14) constituting the battery follows:

> English usage (EU)--a 50-minute test in use of grammar, punctuation word choice, capitalization.
>
> spelling (SP)--a 10-minute test of 50 items, in which the examinee selects the misspelled word from a choice of five words.
>
> reading comprehension (RC)--a 25-minute test of 40 questions on reading paragraphs for stated or implied meanings.
>
> mechanical reasoning (MR)--a 25-minute test of 35 items applying frequently observed mechanical principles and simple physical laws to illustrated problems.
>
> spatial ability (SA)--a 15-minute test of 24 items using spatial reasoning to determine the location of holes in a square paper, given the location of perforations made when the paper is folded.
>
> applied math (AM)--a 20-minute test of 30 items in which arithmetic and elementary algebra are applied to simple practical problems.
>
> vocabulary (VO)--a 25-minute test of 100 antonym items.
>
> data sufficiency (DS)--a 10-minute test of 15 items dealing with elementary mathematics in which the examinee decides whether given statements are sufficient for answering questions.
>
> quantitative judgment (QJ)--a 10-minute test of 30 items in which the examinee determines the relative size of two given quantities.
>
> functional relationships (FR)--a 10-minute test of 15 items in which the examinee determines relationships among two, three, or four variables.
>
> math achievement (MA)--a 60-minute test of 45 items of general mathematics, algebra, and geometry.

These 11 scores along with the three CLEP scores in social science-history (SS), natural science (NS), and humanities (HU) were intercorrelated and then factor analyzed to reveal whether CLEP scores increase the factorial complexity of the WPC battery. Another WPC score, reading speed, which is simply the total number of attempted items on the reading test, was excluded from the analysis for lack of face validity.

Correlations. Means, standard deviations, and intercorrelations between the 14 test scores are shown in Table 1. The natural science CLEP score correlates highly with practically every test in the WPC battery. Only one WPC score, spelling (+.26) correlates less than .40 with natural science, and tests reflecting diverse content such as reading comprehension, mechanical reasoning, vocabulary, quantitative judgment, and math achievement produce relationships greater than .50. Thus the students who performed well on the natural science test after two years of study at the University did quite well on the entire battery of WPC tests taken during the senior year of high school. The social science-history and humanities CLEP scores correlate well with WPC verbal tests, especially reading comprehension and vocabulary, but correlate rather modestly with other tests. Social science-history correlates +.55 with natural science and +.54 with humanities. Humanities correlates +.38 with natural science.

Factor Analysis. The 14 scores were factor analyzed using a computer program described by Kaiser (1959). The matrix of intercorrelations was first analyzed using a principal axes solution. An orthogonal varimax rotation was then carried out to clarify factor labeling and interpretation.

Table 2 displays principal axes factor loadings for the first ten factors, extracted in descending order of explained variance. These factors account for about 92 per cent of the variance in the correlation matrix, with the first factor accounting for better than 44 per cent of the variance. All tests load highly on Factor I, except spelling and CLEP humanities which load moderately. This is apparently a general ability factor, or perhaps a quantitative factor since quantitative tests load somewhat higher than verbal tests. Interestingly, CLEP natural science shows the highest loading on Factor I which is consistent with the earlier observation about its tendency to correlate well with a wide assortment of tests. Factor II, accounting for an additional 15 per cent of the variance, is a verbal factor. Interpretation becomes fuzzy beyond the second factor. This leads to a consideration of the rotated varimax solution results.

The question of how many factors to rotate is seldom easy to resolve (Atkinson and Lunneborg, 1967). The unit latent root criterion would be met by rotating just three factors but these would account for less than two-thirds of the variance. Six factors would be needed to account for more than 80 per cent of the variance, but, through experimentation, it was soon discovered that more than four factors produced instability in results with only one test loading on each additional factor. Four factors (accounting for 72 per cent of the variance) seemed to represent the optimal compromise between maximization of explained variance and substantively meaningful factors. Rotated factor loadings are shown in Table 3.

When rotated, Factor I is clearly a quantitative factor with functional relationships, applied math, math achievement, data sufficiency, and quantitative judgment loading between .65 and .82. Among the three CLEP

scores, natural science loads the highest on this factor, albeit much lower than the five quantitative tests identifying the factor.

Factor II is a verbal factor defined by CLEP humanities, CLEP social science-history, vocabulary, and reading comprehension, with CLEP natural science loading moderately (.55). It is perhaps significant that CLEP natural science, though normally associated with mathematical interests and aptitudes, forms such a strong alliance with the verbal factor.

The third factor, also possessing verbal content, might be labeled grammar or English mechanics. Spelling and English usage are the only tests that load significantly on Factor III.

Mechanical reasoning and spatial ability are the most visible indicators of Factor IV, although CLEP natural science, math achievement, and quantitative judgment also load rather impressively. This pattern of loadings suggests a masculine interest factor. Natural science credits and test achievement are higher for males (Beanblossom, 1969, p. 22) and male superiority in spatial, mechanical, and math achievement areas has been frequently documented.

The factorial results support the contention that CLEP social science-history and CLEP humanities are measuring the usual kinds of verbal abilities found in tests of reading comprehension and vocabulary. Factor II explains about 67 per cent of the variance in humanities (squaring the rotated factor loading) and 63 per cent of the variance in social science-history. Factors I, III, and IV together account for just 5 per cent of the humanities variance and 8 per cent of the social science-history variance. The remaining variance is unexplained.

CLEP natural science is probably the best single indicator, among the 14 tests adopted for this analysis, of what some psychologists have labeled "general intelligence." Masculine interest accounts for about 32 per cent of its variance, the verbal factor 30 per cent, and the quantitative factor 10 per cent. Male students with well-rounded ability profiles are likely to perform well on this exam, but, as with the other two CLEP exams, verbal ability is especially critical.

Conclusions. The evidence presented here plainly shows that the CLEP General Examinations do not measure anything uniquely different from that measured by the traditional pre-college battery of aptitude and achievement tests. This does not mean of course that CLEP tests are not valid indicators of "educational achievement," since students who score well on pre-college exams may very well be more intellectually curious, more highly motivated, and hence better achievers in the liberal arts areas covered by the CLEP exams. But this explanation, by itself, is far from satisfactory since, despite the extreme uneveness of course exposure, and probably interest, in the three CLEP areas, correlations among CLEP tests are quite large. The glow of enthusiasm expressed by test publishers for their own products to impress potential users (psychometric chauvinism?) should be tempered by a disinterested, detached appraisal of the validational claims. In this sample of University juniors the CLEP natural science exam score (probably the most valid of the three exams) is just as predictable from knowledge of a vocabulary or reading comprehension test taken three years ago than a combination of natural science course credits and natural science grades attained in college.

Table 1

Means, Standard Deviations, and Intercorrelations

of WPC and CLEP Scores (N=331)*

	EU	SP	RC	MR	SA	AM	VO	DS	QJ	FR	MA	SS	NS	HU	Mean	S.D.
EU	--	54	50	30	30	38	59	36	38	33	36	36	44	44	51.3	13.6
SP	54	--	34	14	11	28	42	17	20	25	25	23	26	32	22.2	8.9
RC	50	34	--	27	29	37	65	38	36	31	35	55	54	51	16.7	6.3
MR	30	14	27	--	55	51	25	38	53	34	58	30	57	07	12.2	8.1
SA	30	11	29	55	--	48	24	44	49	38	60	23	45	12	12.5	4.5
AM	38	28	37	51	48	--	29	56	64	56	72	33	48	14	15.0	5.1
VO	59	42	65	25	24	29	--	33	37	31	31	56	56	64	63.4	15.0
DS	36	17	38	38	44	56	33	--	54	41	54	27	45	13	8.6	3.1
QJ	38	20	36	53	49	64	37	54	--	49	71	34	55	20	18.2	5.3
FR	33	25	31	34	38	56	31	41	49	--	59	29	40	19	7.5	4.3
MA	36	25	35	58	60	72	31	54	71	59	--	36	59	14	26.9	9.9
SS	36	23	55	30	23	33	56	27	34	29	36	--	55	54	48.3	12.6
NS	44	26	5	57	45	48	56	45	55	40	59	55	--	38	53.8	16.8
HU	44	32	51	07	12	14	64	13	20	19	14	54	38	--	33.1	13.3

*Decimal points omitted.

Table 2

Fourteen Variable Principal Axes Factor Loading Matrix

	Factor									
Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
English Usage	.662	.337	-.375	-.223	-.149	-.087	-.161	-.127	.047	-.369
Spelling	.455	.363	-.665	-.219	.168	.150	.158	.157	.084	.229
Reading Comprehension	.685	.407	.125	.081	-.214	.025	.334	-.056	-.387	-.020
Mechanical Reasoning	.641	-.378	.194	-.448	.146	.130	.019	-.188	.059	-.124
Spatial Ability	.622	-.383	.131	-.311	-.093	-.502	.149	.204	.020	.062
Applied Math	.745	-.352	-.168	.167	.044	.142	.049	.202	-.147	-.187
Vocabulary	.689	.536	.061	-.006	-.091	-.057	-.104	-.167	-.051	.106
Data Sufficiency	.654	-.259	-.060	.223	-.574	.100	.030	.008	.303	.073
Quantitative Judgment	.754	-.313	.003	.037	.009	.124	-.405	.084	-.196	.074
Functional Relationships	.636	-.227	-.213	.448	.312	-.264	.112	-.317	.091	-.013
Math Achievement	.787	-.410	-.043	.033	.137	.008	-.030	.127	-.057	.081
CLEP Social Science-History	.628	.365	.402	.133	.215	.189	.155	.218	.238	-.188
CLEP Natural Science	.792	.032	.270	-.156	.057	.177	-.002	-.225	.058	.242
CLEP Humanities	.488	.664	.170	.036	.092	-.247	-.248	.169	.063	.057
Latent root	6.22	2.07	1.00	.74	.64	.55	.46	.44	.39	.37
% Variance	44.43	14.79	7.31	5.31	4.57	3.90	3.30	3.14	2.81	2.64
Cumulative % Variance	44.43	59.22	66.35	71.66	76.23	80.13	83.43	86.57	89.38	92.02

Table 3

Fourteen Variable Varimax Rotated Factor Loading Matrix

	Factor				
Variable	I	II	III	IV	Communality
English Usage	.205	.394	(.705)	.218	.7418
Spelling	.139	.173	(.883)	.010	.8288
Reading Comp.	.250	(.723)	.223	.147	.6569
Mechanical Reasoning	.246	.098	.073	(.846)	.7907
Spatial Ability	.335	.079	.066	(.725)	.6479
Applied Math	(.755)	.110	.180	.347	.7340
Vocabulary	.157	(.774)	.353	.134	.7665
Data Sufficiency	(.666)	.182	.077	.257	.5484
Quant. Judgment	(.654)	.204	.088	.444	.6741
Func. Relationships	(.812)	.163	.124	.031	.7017
Math Achievement	(.715)	.130	.110	(.504)	.7939
CLEP SocSci-History	.199	(.792)	-.051	.189	.7058
CLEP Natural Science	.309	(.546)	.095	(.568)	.7249
CLEP Humanities	.011	(.818)	.211	-.056	.7169

References

Atkinson, G. and Lunneborg, C. E. Comparison of oblique and orthogonal simple structure solutions for personality and interest factors. Seattle: Bureau of Testing, University of Washington, 1967. (Duplicated report)

Beanblossom, G. F. The use of CLEP scores in evaluating liberal arts curriculum. Seattle: Bureau of Testing, University of Washington, 1969. (Duplicated report)

College-Level Examination Program. A description of the general examinations. Princeton, New Jersey: College Entrance Examination Board, 1967.

College-Level Examination Program. Description and uses, 1968. Princeton, New Jersey: College Entrance Examination Board, 1968.

Kaiser, H. F. Computer program for varimax rotation in factor analysis. Educational and Psychological Measurement, 1959, 19, 413-420.

Washington Pre-College Testing Program. Counselor's manual 1968-69. Seattle: WPCTP, 1968.